

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-152894

(43) 公開日 平成9年(1997)6月10日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/00			G 1 0 L 9/00	C
3/00	5 1 3		3/00	5 1 3 A

審査請求 未請求 請求項の数 6 O L (全 13 頁)

(21) 出願番号 特願平7-312814

(22) 出願日 平成7年(1995)11月30日

(71) 出願人 000004260

株式会社デンソー

愛知県刈谷市昭和町1丁目1番地

(72) 発明者 中村 一雄

愛知県刈谷市昭和町1丁目1番地 日本電

装株式会社内

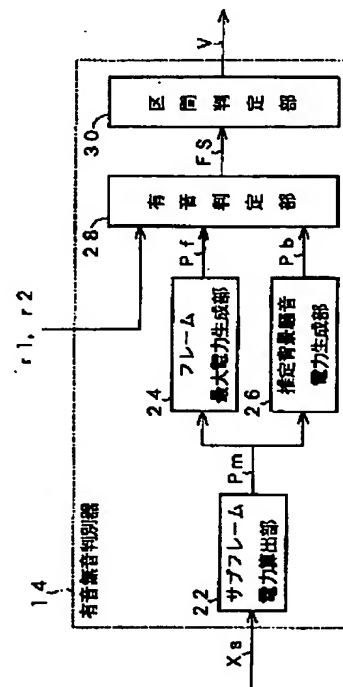
(74) 代理人 弁理士 足立 勉

(54) 【発明の名称】 有音無音判別器

(57) 【要約】

【課題】 音声の語頭部分を含むフレームの有音無音を正確に判定でき、背景騒音が大きい等、悪環境においても誤判定の少ない有音無音判別器を提供する。

【解決手段】 サブフレーム電力算出部22では、フレームを4分割したサブフレーム毎にサブフレーム電力 $P_m$ を算出し、このサブフレーム電力 $P_m$ に基づいて、フレーム最大電力生成部24では、サブフレーム毎に一つのサブフレーム電力との移動平均（短期平均値）を算出すると共に、同一フレームを構成するサブフレーム間で短期平均値を比較し最大のものを該フレームのフレーム最大電力 $P_f$ として選択する。これにより、発声がフレームの後半から開始されたとしても、フレーム最大電力 $P_f$ が小さく見積られることがなく、該フレームは有音判定部28にて、確実に有音と判定される。



## 【特許請求の範囲】

【請求項 1】 入力音声信号を所定時間毎の基本フレーム単位に分割し、該基本フレーム毎に有音無音を判別する有音無音判別器であって、

上記音声信号が、上記基本フレームを更に所定分割数に分割したサブフレーム分だけ入力される毎に、該サブフレームの電力値であるサブフレーム電力を順次算出するサブフレーム電力算出手段と、

上記音声信号が、上記基本フレーム分だけ入力される毎に、上記サブフレーム電力算出手段にて算出された該基本フレームを構成するサブフレームの各サブフレーム電力のうち、最大のものを該基本フレームのフレーム最大電力とするフレーム最大電力生成手段と、

上記サブフレーム電力算出手段にて算出された最新のサブフレーム電力を含む、該サブフレーム電力以前に算出された連続する複数のサブフレーム電力に基づき、背景騒音電力を推定する背景騒音電力推定手段と、

上記フレーム最大電力と上記背景騒音電力との差に基づき、上記基本フレームの有音無音を判別する有音無音判別手段と、

を備えることを特徴とする有音無音判別器。

【請求項 2】 上記フレーム最大電力生成手段は、上記サブフレーム算出手段にてサブフレーム電力が算出される毎に、該最新のサブフレーム電力を含む、該サブフレーム電力以前に算出された上記分割数より少ない所定個の連続するサブフレーム電力に基づき、その平均値である短期平均値を算出する短期平均値算出手段を備え、

上記サブフレーム電力に代えて上記短期平均値が最大のものをフレーム最大電力とすることを特徴とする請求項 1 に記載の有音無音判別器。

【請求項 3】 上記背景騒音電力推定手段は、上記サブフレーム電力算出手段にてサブフレーム電力が算出される毎に、該最新のサブフレーム電力を含む、該サブフレーム電力以前に算出された上記分割数より多い所定個の連続するサブフレーム電力に基づき、その平均値である長期平均値を算出する長期平均値算出手段と、上記音声信号が、上記基本フレーム分だけ入力される毎に、上記長期平均値算出手段にて算出された該基本フレームを構成するサブフレームの各長期平均値のうち、最小のものを該基本フレームの背景騒音電力とする選択手段と、

により構成されていることを特徴とする請求項 1 または請求項 2 に記載の有音無音判別器。

【請求項 4】 請求項 1 ないし請求項 3 のいずれかに記載の有音無音判別器において、更に、上記入力音声信号を上記基本フレーム毎に線形予測分析して、該入力音声信号の周波数スペクトル包絡線の特徴を表す特徴パラメータを抽出するパラメータ抽出手段を備え、

上記有音無音判別手段は、

上記フレーム最大電力と上記背景騒音電力との電力差が所定の第 1 しきい値以上である基本フレームを有音、上記電力差が上記第 1 しきい値より小さい所定の第 2 しきい値以下である基本フレームを無音と判定する第 1 の判定手段と、

上記電力差が第 1 しきい値より大きく、且つ第 2 しきい値より小さい場合に、上記パラメータ抽出手段にて抽出される特徴パラメータに基づき、上記基本フレームの有音無音を判定する第 2 の判定手段と、

により構成されていることを特徴とする有音無音判別器。

【請求項 5】 請求項 4 に記載の有音無音判別器において、

上記パラメータ抽出手段にて抽出され、上記第 2 の判定手段にて判定に使用される特徴パラメータは、低次の反射係数であることを特徴とする有音無音判別器。

【請求項 6】 請求項 1 ないし請求項 5 のいずれかに記載の有音無音判別器において、

更に、上記有音無音判別手段にて有音と判定された有音フレーム及び無音と判定された無音フレームのうち、有音フレームと該有音フレームの後に連続する所定個以内の無音フレームとを有音区間と判定し、有音フレームの後に連続する所定個の無音フレームの後に、更に連続する無音フレームを無音区間と判定する区間判定手段を備えたことを特徴とする有音無音判別器。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、音声分析合成系において入力音声の有音無音を判別する有音無音判別器に関し、特に、一定時間毎のフレーム単位で音声信号を分析して得られる符号化データを送受信するデジタル移動通信において、移動局の連続使用時間を増大させる目的で行う低消費電力制御のひとつである V O X (Voice Operated Transmitter) 制御に適用することを目的とした有音無音判別器に関する。

## 【0002】

【従来の技術】従来より、移動体通信網で使用される移動局は、電源である電池容量の制約から連続通話時間が有限であるため、様々な低消費電力制御を施して連続通話時間の延長を図る努力がなされている。

【0003】その中の一つの技術として、通話中の会話音声の出現頻度が小さいことを利用し、相手の話を聞いている場合等、音声を発していない時には、送信回路を停止することにより消費電力を抑え、通話時間の延長を図る V O X (Voice Operated Transmitter) 制御が知られている。

【0004】そして、このような V O X 制御を実現するためには、送信する音声の有音無音を正確に判定することが要求されるが、この種の装置として、例えば、特開

## 3

平5-323996号公報には、音声信号を所定時間毎のフレーム単位で高能率符号化する音声符号化器を備えた送信器において使用され、符号化を行うフレーム単位で、音声の有音無音を判定する有音無音判定器が開示されている。

【0005】この有音無音判定器では、音声の送信を行っている有音区間では、各フレーム毎に求めたフレームの平均電力と、音声符号化器において音声信号を高能率符号化する過程で求められ音声の特徴を表すパラメータとを用いてフレームの有音無音を判定し、一方、音声の送信を停止している無音区間では、電力消費の低減のために音声符号化器も停止させているため、フレームの平均電力のみに基づいてフレームの有音無音を判定するように構成されている。

## 【0006】

【発明が解決しようとする課題】しかし、この装置では、無音区間でのフレームの有音無音の判定を、フレームの平均電力のみによっているため、発声が始まった時に、語頭切れが発生しやすく、音声の自然性が損なわれてしまうという問題があった。

【0007】即ち、この種の装置において用いられるフレームのサイズは、通常20ms程度と長いため、フレームの後半から発声が始まった場合、音声の語頭部分を含むフレームの平均電力は小さなものとなるため、無音フレームと誤判定されやすく、その結果、語頭切れが生じてしまうのである。

【0008】また、この語頭切れを防止するために、フレームの有音無音を判定するしきい値を小さな値に設定すると、背景騒音の変動等の影響を受け易くなり、無音フレームを有音フレームであると判定する誤動作が頻発するようになるため、VOX制御を有効に実施することができないという問題もあった。

【0009】本発明は、上記問題点を解決するために、音声の語頭部分を含むフレームの有音無音を正確に判定でき、しかも背景騒音が大きい等、悪環境においても誤判定の少ない有音無音判別器を提供することを目的とする。

## 【0010】

【課題を解決するための手段】上記目的を達成するためになされた請求項1に記載の発明においては、基本フレームを所定分割数に分割してなるサブフレーム分だけ音声信号が入力されると、サブフレーム電力算出手段が、このサブフレームの電力値であるサブフレーム電力を算出し、その後、基本フレーム分だけ音声信号が入力されると、フレーム最大電力生成手段が、その基本フレームを構成するサブフレームの各サブフレーム電力のうち、最大のものを該基本フレームのフレーム最大電力として選択する。

【0011】一方、背景騒音電力推定手段では、サブフレーム電力算出手段にて算出された最新のサブフレーム

## 4

電力を含む、該サブフレーム電力以前に算出された連続する複数のサブフレーム電力に基づき、背景騒音電力を推定する。そして、有音無音判別手段が、基本フレーム毎に、フレーム最大電力と背景騒音電力との差に基づいて、その有音無音を判別する。

【0012】このように、本発明においては、サブフレーム毎にサブフレーム電力を算出し、同一基本フレームを構成するサブフレーム間でこれを比較し、その最大値をフレーム最大電力として選択して、基本フレームの有音無音の判定に使用するようにされている。

【0013】従って、本発明の有音無音判定器によれば、基本フレームの後半から発声が始まったとしても、その音声信号の電力が基本フレーム全体で平均化されてしまうことがなく、基本フレーム全体に音声信号が存在する場合と同様に、ほぼ音声信号の電力レベルに即したフレーム最大電力が得られるため、確実にこれを有音フレームと判定することができる。その結果、当該有音無音判定器を、音声の基本フレーム単位で符号化して送信する送信器に適用した場合に、音声の語頭切れを確実に防止でき、音声の品質を向上させることができる。

【0014】また、語頭切れ防止のために、フレーム最大電力と背景騒音電力との差に基づいて有音無音を判定するためのしきい値を低く設定する必要がないため、無音を有音と判定する誤判定も確実に減少させることができ、有効なVOX制御を実現することができる。

【0015】また、従来装置のように、基本フレーム単位で処理を行っている場合、基本フレームが大きい程、有音無音の判定精度は劣化し、語頭切れや誤動作が発生し易いが、本発明によれば、サブフレームの大きさにより、判定精度が決まるため、基本フレームの大きさに関係なく、判定精度を任意に設定することができる。

【0016】次に、請求項2に記載の発明においては、フレーム最大電力生成手段は、短期平均値算出手段を備え、この短期平均値算出手段は、サブフレーム算出手段にてサブフレーム電力が算出される毎に、該最新のサブフレーム電力を含む、該サブフレーム電力以前に算出された分割数より少ない所定個の連続するサブフレーム電力に基づき、その平均値である短期平均値を算出する。

【0017】そして、フレーム最大電力生成手段は、サブフレーム電力に代えて、この短期平均値に基づき、フレーム最大電力を選択する。つまり、本発明においては、突発的に振幅の大きなノイズが音声信号に重畳されると、その影響を直接受けてしまうサブフレーム電力の代わりに、サブフレーム電力の移動平均である短期平均値を使用してフレーム最大電力を求めることにより、このようなノイズの影響を緩和するようにされている。

【0018】従って、本発明の有音無音判別器によれば、突発的なノイズ等の影響による誤判定を確実に低減することができる。次に、請求項3に記載の発明においては、背景騒音電力推定手段が、長期平均値算出手段と

10

20

30

40

50

## 5

選択手段とからなり、長期平均値算出手段が、サブフレーム電力算出手段にてサブフレーム電力が算出されると、この最新のサブフレーム電力を含み、このサブフレーム電力以前に算出された分割数より多い所定個の連続するサブフレーム電力に基づき、その平均値である長期平均値を順次算出し、その後、基本フレーム分だけ音声信号が入力されると、選択手段が、長期平均値算出手段にて算出された同一基本フレームを構成するサブフレームの各長期平均値のうち、最小のものを該基本フレームの背景騒音電力として選択する。

【0019】つまり、本発明においては、背景雑音電力を複数基本フレームに渡るサブフレーム電力の移動平均に基づいて求めている。従って、本発明の有音無音判別器によれば、基本フレームの有音無音を判定する際の基準レベルとなる背景騒音電力は、突発的なサブフレーム電力の変動に追従して急激に変化してしまうことがないため、基本フレームの有音無音の判定を安定して行うことができる。

【0020】次に、請求項4に記載の有音無音判別器においては、パラメータ抽出手段が、入力音声信号を上記基本フレーム毎に線形予測分析して、該入力音声信号の周波数スペクトル包絡線の特徴を表す特徴パラメータを抽出し、有音無音判別手段では、第1の判定手段が、フレーム最大電力と背景騒音電力との電力差が所定の第1しきい値以上である基本フレームを有音、また、電力差が第1しきい値より小さい所定の第2しきい値以下である基本フレームを無音と判定し、電力差が第1しきい値より大きく、且つ第2しきい値より小さい場合には、第2の判定手段が、パラメータ抽出手段にて抽出される特徴パラメータに基づき、基本フレームの有音無音を判定する。

【0021】ところで、一般に、線形予測分析により抽出される特徴パラメータは、音声信号の周波数スペクトル包絡線の特徴を表すものであり、有音時と無音時とでは夫々互いに異なった特徴的な傾向を示すため、この特

$$r_n = -\frac{R_n}{R_0}$$

【0027】なお、 $n$ 次の自己相関 $R_n$ とは、(2)式に示すように、離散的に表された音声波形のデータ $x_i$ を、 $n$ 個前のデータ $x_{i-n}$ と掛け合わせて、それを全領

$$R_n = \sum_i \{x_i \cdot x_{i-n}\}$$

【0029】そして、低次の反射係数は、その分布が有声音（母音等）と無声音（無声子音や背景騒音等）とで比較的良好に分離していることが知られている。例えば、1次の反射係数 $r_1$ の場合を考えると、有声音は、低周波領域でははっきりしたホルマント構造を持つので、その一次相関 $R_1$ は、全エネルギー $R_0$ と類似の値を示し、無声音は、周波数スペクトルの偏りを示さないことが多いため、その一時相関 $R_1$ は、小さい値を示す。その結

## 6

微パラメータを用いることにより、背景騒音電力の大きさによらず、有音と無音とをある程度判別し得ることが知られている。

【0022】そして、本発明においては、フレーム最大電力と背景騒音電力との差に基づいて、有音無音を判定する場合に、それ以上であれば、ほぼ確実に有音であると判定できるようなレベルに第1しきい値を設定し、それ以下であれば、ほぼ確実に無音であると判定できるようなレベルに第2しきい値を設定し、フレーム最大電力と背景騒音電力との差では、誤判定を起こす可能性の高いこれらの間の領域において、特徴パラメータによる判定を行うようにされている。

【0023】従って、本発明の有音無音判別器によれば、背景騒音電力が大きく、フレーム最大電力と背景騒音電力との差が小さいときであっても、より確実に有音無音を判定することができ、単に、フレーム最大電力と背景騒音電力との差に基づいて判定を行う場合に比べて、有音無音の判定精度を向上させることができる。

【0024】即ち、特徴パラメータは、基本フレームの有音無音を確実に判定できるものではないが、背景騒音電力の大きさによらず、判定が可能であるため、フレーム最大電力により判定を下すのが微妙な範囲において、この特徴パラメータによる判定を補助的に使用することにより、有音無音の判定精度を効果的に向上させることができるのである。

【0025】次に請求項5に記載の有音無音判別器においては、パラメータ抽出手段にて抽出され、第2の判定手段にて判定に使用される特徴パラメータは、低次の反射係数であることを特徴とする。ここで、 $n$ 次の反射係数 $r_n$ とは、(1)式に示すように、入力信号の $n$ 次の自己相関 $R_n$ を全エネルギー（0次の自己相関） $R_0$ で正規化したものに相当する。

【0026】

【数1】

(1)

域について加算したものである。

【0028】

【数2】

(2)

果、1次の反射係数 $r_1$ は、有声音の場合、ほぼ1となり、無声音の場合、ほぼ0となる。

【0030】更に、無声音には、無声子音と背景騒音とが含まれるが、これらの周波数スペクトルの概形の差も1次の反射係数 $r_1$ に反映され、即ち、周波数スペクトルが3～10kHzの高周波領域に比較的大きな成分を有し高域強調の特性となる無声子音では、1次の反射係数 $r_1$ は+1側に偏り、一方、周波数特性が-9dB/

## 7

oct 程度の傾斜を有し低域強調の特性となる背景騒音では、1次の反射係数 $r_1$ は-1側に偏る。

【0031】このように、1次の反射係数 $r_1$ は、入力音声信号が有音の場合、即ち有声音や無声子音を含んでいる場合には+1に近い値となり、入力音声信号が無音の場合、即ち背景騒音のみの場合には-1に近い値となっており、よく分離された特徴を示す。

【0032】なお、母音などの有声音も周囲騒音と同様に低域強調の周波数特性を示すが、通常、背景騒音に比べて十分に大きなパワーを有しているため、通常は、フレーム最大電力と背景騒音電力との差に基づく有音無音判定で、充分に分離される。従って、このような特徴を備えた低次の反射係数 $r_n$ による有音無音判定を、フレーム最大電力と背景騒音電力との差に基づく有音無音判定と併用することにより、基本フレームの有音無音の判定精度を、より向上させることができる。

【0033】次に、請求項6に記載の有音無音判別器においては、区間判定手段が、有音無音判別手段にて有音と判定された有音フレーム及び無音と判定された無音フレームのうち、有音フレームと該有音フレームの後に連続する所定個以内の無音フレームとを有音区間と判定し、有音フレームの後に連続する所定個の無音フレームの後に、更に連続する無音フレームを無音区間と判定する。

【0034】つまり、有音フレームの直後に無音フレームが表れても、すぐにこれを無音区間と判定することはせず、無音フレームが所定個以上連続した場合に、それ以降の無音フレームを無音区間と判定するようにされており、会話中の息継ぎ等、発声時における短時間の無音が、いちいち無音区間として判定される等して、頻繁に有音区間と無音区間とが切り替わってしまうことがないようにされている。

【0035】従って、本発明の有音無音判別器によれば、ここで判定される無音区間／有音区間に基づいてV O X制御を行う場合に、V O X制御が頻繁に行われることによる語頭切れ等による違和感を防止できる。また、有音フレーム直後の無音フレームが欠落することがないので、音声の語尾切れを確実に防止できる。

## 【0036】

【発明の実施の形態】以下に本発明の実施例を図面と共に説明する。図2は、本実施例の有音無音判別器が適用された、デジタル方式の自動車電話システムにて用いられる移動局の送信部の構成を表すブロック図である。

【0037】図2に示すように、移動局の送信部は、マイクを介して入力され電気信号に変換された音声信号を125 $\mu$ s周期(8kHz)でサンプリングして、16ビットのデジタルデータXsに変換する入力部10と、入力部10からのデジタルデータXsを、20ms(160データ)毎に区切ったフレーム単位で分析処理して高能率符号化する音声符号化器12と、入力部10から

## 8

入力されるデジタルデータXsに基づき、音声符号化器12にて処理されたフレームが、音声信号を含む有音フレームであるか音声信号を含まない無音フレームであるかを判定し、その判定に基づき、該フレームが有音区間／無音区間のいずれに属するかを表す判定結果Vを出力する有音無音判別器14と、音声符号化器12により生成された符号化データの送信を、有音無音判別器14からの判定結果Vに基づき制御する送信制御部16と、送信制御部16からの制御に従って所定フォーマットの送信フレームを生成し、アンテナ20を介して送信する送信器18と、により構成されている。

【0038】ここで、音声符号化器12は、デジタルシグナルプロセッサ(DSP)により構成されており、デジタル方式の自動車電話システムの標準規格として定められた音声符号化方式である周知のVSELP(Vector Sum Excited Linear Prediction)方式により音声データを符号化するものである。

【0039】VSELP方式の音声符号化器12では、音声の高能率符号化技術の基本モデルである声帯(音源)及び声道からなる音声生成モデルに基づいて、声帯から発せられる音を模擬した各種音源情報、声道を模擬した声道等価フィルタを構成するためのフィルタ係数等を符号化データとして出力する。

【0040】このうち、フィルタ係数は、所定のフレーム単位で音声信号を線形予測分析することにより求められる。なお、このフィルタ係数の算出過程において1次から10次の反射係数が求められ、そのうち1次と2次の反射係数 $r_1$ 、 $r_2$ が、有音無音判別器14に入力される。

【0041】また、音源情報としては、声帯から発せられる音を模擬した音源の駆動信号をベクトル量子化し、これをコードベクトルと呼んで、複数用意されたコードベクトルの中から、合成音声のひずみが最小となるようなものが選択され出力されるのであるが、このコードベクトルが直接送信されるのではなく、これを間接的に指定する基底ベクトルのインデックスが用いられる。また、音源情報としては、フレームの電力値等も出力される。

【0042】次に、送信制御部16は、有音無音判別器14による判定結果Vが無音区間( $V=0$ )から有音区間( $V=1$ )に変化すると、送信器18を起動し、送信開始を予告するための信号として2フレーム分(40ms)のプリアンプル生成して、送信器18に入力する。引続き、判定結果Vが有音区間( $V=1$ )である限り、音声符号化器12からの符号化データを連続して送信器18に転送する。その後、判定結果Vが有音区間( $V=1$ )から無音区間( $V=0$ )に変化すると、未送信の符号化データ(ここでは2フレーム分)があれば、これを送信器18に転送し、更に、送信停止を予告するための信号として3フレーム分(60ms)のポストアンプル

生成して送信器18に入力後、送信器18を停止する。

【0043】また、判定結果Vとして無音区間（V＝0）が長時間継続している場合には、50フレーム（1000ms）毎に送信器18を起動し、ポストアンプを生成して送信器18に入力してこれを送信させた後、再度送信器18を停止する。なお、送信器18は、送信制御部16からのプリアンプ、符号化データ、パストアンプを、所定の送信符号に変換する等の処理を施した上で送信する。図3は、このとき、送信器18から送信される送信フレームの構成を表す説明図である。

【0044】次に、有音無音判別器14は、図1に示すように、入力部10からのデジタルデータXsを順次入力し、音声符号化器12にて符号化される単位のフレームを、更に4分割して5ms（40データ）毎に区切ったサブフレーム単位で、そのサブフレーム内の電力Pmを算出するサブフレーム電力算出部22と、サブフレーム電力算出部22にて順次算出されるサブフレーム電力Pmに基づき、フレーム最大電力Pfを生成するフレーム最大電力生成部24と、同じくサブフレーム電力Pmに基づき推定背景騒音電力Pbを生成する推定背景騒音電力生成部26と、フレーム最大電力Pf、推定背景騒音電力Pb、音声符号化器12からの1次及び2次の反射係数r1、r2に基づき、フレームの有音無音を判定する有音判定部28と、有音判定部28の判定結果であるフレーム状態Fsに基づき、符号化データの送信を行うべき有音区間と、送信を停止すべき無音区間とを判定する区間判定部30と、により構成されている。

【0045】なお、有音無音判別器14は、CPU、ROM、RAMを中心に構成された周知のマイクロプロセッサからなり、本実施例においては、音声符号化器12（DSP）と有音無音判別器14とは同一のワンチップのマイクロプロセッサから構成されている。また、有音無音判別器14を構成している上記サブフレーム電力算出部22、フレーム最大電力生成部24、推定背景騒音電力生成部26、有音判定部28、区間判定部30は、CPU上で実行される処理として実現されている。

【0046】ここで、有音無音判別器14の各部の処理をフローチャートに沿って説明する。図4は、サブフレーム電力算出部22に相当するサブフレーム電力算出処理を表すフローチャートである。本処理は、入力部10にて入力音声信号がサンプリングされる毎、即ち125μs周期で起動される。

【0047】なお、本処理にて使用される各変数Ps、

$$PAL(i, j) = \frac{\sum_{k=0}^{Nb-1} Pm(k)}{Nb}$$

【0053】

nは、当該有音無音判別器14への電源投入直後に実行される初期化処理にて0クリアされているものとする。本処理が起動されると、まず、ステップ110では、入力部10からのデジタルデータXs<sup>2</sup>を変数Psに加算することにより、デジタルデータXs<sup>2</sup>の累積値、即ち、信号電力の積分値を求め、続くステップ120では、デジタルデータXs<sup>2</sup>の加算回数を表すカウント値nをインクリメントしてステップ130に進む。

【0048】ステップ130では、カウント値nが、1サブフレームに相当する値Nsmpl（本実施例では、40）より小さいか否かを判断し、小さければそのまま本処理を終了する。一方、カウント値nが比較値Nsmplに達している場合には、1サブフレームの電力の算出が終了したものとしてステップ140に移行し、デジタルデータXs<sup>2</sup>の累積値、即ちサブフレーム電力Psを所定のバッファに格納する。

【0049】なお、バッファは、リングバッファ形式で構成されており、常に、最新のサブフレーム電力Psが、後述する長期平均値を算出するのに必要な所定個Nbuffだけ記憶されるようにされている。続くステップ150では、次のサブフレームの処理に備えるために累積値Ps、カウント値nを0クリアし、続くステップ160では、フレーム最大電力生成部24、推定背景騒音電力生成部26に相当する背景騒音電力、フレーム最大電力算出処理を起動して本処理を終了する。

【0050】次に、図5は、背景騒音電力、フレーム最大電力算出処理を表すフローチャートである。本処理は、先のサブフレーム電力算出処理により、1サブフレームのサブフレーム電力の算出が終了する毎、即ちサブフレーム間隔（5ms）で周期的に起動される。

【0051】なお、本処理において使用される変数i、jは、現在処理中のフレームi、及びサブフレームjを識別するための変数であり、先のサブフレーム電力算出処理における変数と同様に、初期化処理にて0クリアされているものとする。図5に示すように、本処理が起動されると、まず、ステップ210にて、（2）式に基づき、サブフレーム電力の長期平均値PAL(i, j)が算出され、続くステップ220では、（3）式に基づき、サブフレーム電力の短期平均値PAS(i, j)が算出される。

【0052】

【数3】

(3)

【数4】

$$PAS(i, j) = \frac{\sum_{k=0}^{Nf-1} Pm(k)}{Nf}$$

(4)

【0054】なお、 $Pm(k)$ は、先のステップ140にて、バッファに格納されたサブフレーム電力値であり、今回算出された最新のサブフレーム電力値を $Pm(0)$ 、 $k$ 回前に算出されたサブフレーム電力値を $Pm(k)$ とする。

【0055】即ち、サブフレーム電力の長期平均値 $PAL(i, j)$ とは、図6に示すように、サブフレーム電力算出処理にて算出された最新のサブフレーム電力値を含む、この最新のサブフレーム以前の算出された連続する所定個 $Nb$ （本実施例では16個）のサブフレーム電力値の平均値のことであり、一方、サブフレーム電力の短期平均値 $PAS(i, j)$ とは、最新のサブフレーム電力値を含む、この最新のサブフレーム電力値以前に算出された連続する所定個 $Nf$ 個（本実施例では2個）のサブフレーム電力値の平均値のことであり、いずれも、サブフレーム電力値の移動平均を算出するものである。

【0056】続くステップ230では、変数 $j$ をインクリメントしてステップ240に進む。ステップ240では、変数 $j$ が、1フレーム内のサブフレーム数 $Nsf$ （ここでは4）より小さいか否かを判断し、小さければ1フレーム分の処理はまだ終了していないものとして、本処理を終了する。一方、変数 $j$ がサブフレーム数 $Nsf$ に達していれば、1フレーム分の処理が終了したものとして、ステップ250に移行し、(4)式に基づき、当該フレームを構成する各サブフレーム毎に算出された長期平均値 $PAL(i, j)$ のうち、最小のものを当該フレームにおける背景騒音電力 $Pb(i)$ として算出し、続くステップ260では、(5)式に基づき、短期平均値 $PAS(i, j)$ のうち、最大のものを当該フレームにおけるフレーム最大電力 $Pf(i)$ として算出する。

【0057】

20 【数5】

$$Pb(i) = \min\{PAL(i, 0), PAL(i, 1), PAL(i, 2), PAL(i, 3)\} \quad (5)$$

【0058】

【数6】

$$Pf(i) = \max\{PAS(i, 0), PAS(i, 1), PAS(i, 2), PAS(i, 3)\} \quad (6)$$

【0059】続くステップ270では、フレームを識別する変数 $i$ をインクリメントすると共に、サブフレームを識別する変数 $j$ を0クリアする。続くステップ280では、有音判定部28に相当する有音無音判定処理を起動して、本処理を終了する。

【0060】次に、図7は、有音無音判定処理を表すフローチャートである。本処理は、背景騒音電力、フレーム最大電力算出処理により、1フレームの背景騒音電力、フレーム最大電力が算出される毎、即ちフレーム間隔（20ms）で周期的に起動される。

【0061】図7に示すように、本処理が起動されると、まずステップ310にて、フレーム最大電力 $Pf(i)$ が、背景騒音電力 $Pb(i)$ に第1しきい値 $TH1$ （本実施例では50）を加えた比較値以上であるか否かを判断し、比較値以上であればステップ360に移行し、比較値より小さければステップ320に移行する。

【0062】ステップ320では、フレーム最大電力 $Pf(i)$ が、背景騒音電力 $Pb(i)$ に第2しきい値 $TH2$ （本実施例では30）を加えた比較値以上であるか否かを判断し、比較値以上であればステップ330に移行し、比較値より小さければステップ350に移行する。

【0063】ステップ330では、音声符号化器12にて、 $i$ フレームを線形予測分析することにより算出された1次の反射係数 $r1(i)$ が、第3しきい値 $TH3$

（本実施例では0.7）以上であるか否かを判断し、第3しきい値 $TH3$ 以上であれば、ステップ360に移行し、第3しきい値 $TH3$ より小さければステップ340

に移行する。

【0064】ステップ340では、2次の反射係数 $r2(i)$ が、第4しきい値 $TH4$ 以下であるか否かを判断し、第4しきい値 $TH4$ （本実施例では0.3）以下であればステップ360に移行し、第4しきい値 $TH4$ より大きければステップ350に移行する。

【0065】そして、ステップ350では、フレームの有音無音の判定結果を表す変数 $FS$ を無音を表す値（ $FS=0$ ）に設定してステップ370に進み、一方、ステップ360では、変数 $FS$ を有音を表す値（ $FS=1$ ）に設定してステップ370に進む。

【0066】ステップ370では、区間判定部30に相当する区間判定処理を実行して、本処理を終了する。即ち、フレーム最大電力 $Pf(i)$ が、背景騒音電力 $Pb(i)$ に、第1しきい値 $TH1$ を加算した値以上である場合には、このフレームを有音であると判定し、フレーム最大電力 $Pf(i)$ が、背景騒音電力 $Pb(i)$ に第2しきい値 $TH2$ （ $<TH1$ ）を加算した値より小さい場合には、このフレームを無音であると判定する。そして、フレーム最大電力 $Pf(i)$ が、これらの中間の大きさである時には、1次の反射係数 $r1(i)$ が第3しきい値 $TH3$ 以上であるか、2次の反射係数 $r2(i)$ が第4しきい値 $TH4$ 以下である場合に、このフレームを有音であると判定し、それ以外の場合を無音であると判断する。

【0067】なお、各値は16ビットで表されるものとする。次に、区間判定処理の処理を、図8に示す状態遷移図を用いて説明する。なお、本処理は、先の有音無音



判定処理により設定されたフレーム状態  $FS$  に基づき、有音状態、保留状態、無音状態の3状態の間で状態を遷移させ、これら状態に応じて、有音区間／無音区間を表す判定結果  $V$  を設定するものである。

【0068】まず、有音状態とは、前回処理されたフレームが有音 ( $FS=1$ ) であった場合であり、この状態では、判定結果  $V$  は有音区間を表す値 ( $V=1$ ) に設定されている。そして、新たに処理されたフレームが、有音無音判定処理にて有音 ( $FS=1$ ) と判定された場合は、そのまま有音状態に留まる。一方、新たに処理されたフレームが、無音 ( $FS=0$ ) と判定された場合は、保留状態に遷移すると共に、無音フレームカウンタ  $C$  に、所定値  $N_{wait}$  (本実施例では、20) を設定する。なお、この保留状態では、判定結果  $V$  は、有音状態と同様に有音区間を表す値 ( $V=1$ ) に保持される。

【0069】そして、保留状態では、新たに処理されたフレームが有音と判定された場合は、有音状態に遷移する。一方、無音と判定された場合は、無音フレームカウンタ  $C$  をデクリメントし、その値が0でない時には、そのまま保留状態に留まり、0の時には、判定結果  $V$  を無音区間を表す値 ( $V=0$ ) に設定して無音状態に遷移する。

【0070】この無音状態では、新たに処理されたフレームが無音と判定された場合は、そのまま無音状態に留まり、一方、有音と判定された場合は、判定結果  $V$  を有音区間を表す値 ( $V=1$ ) に設定して有音状態に遷移する。即ち、本処理においては、有音状態及び保留状態が有音区間 ( $V=1$ )、無音状態が無音区間 ( $V=0$ ) と判定され、これに応じて判定結果  $V$  が設定される。

【0071】そして、無音区間にある時には、有音フレームを一つ受信すると、直ちに有音区間と判定され、一方、有音区間にある時には、無音フレームを20個以上連続して受信した場合、即ち、無音が400ms以上継続した場合に無音区間と判定される。

【0072】このようにして、判定された有音区間／無音区間の判定結果  $V$  が、送信制御部16に入力されることにより、先に説明したように、送信器18の起動／停止、及び符号化データの送信が制御される。ここで、図9、10、11は、静かな室内で、約2秒間隔で発声した音声に、走行中の車両の車室内で録音した騒音を  $S/N$  比15dBで重畳したものを入力音声として、シミュレーションを行った結果を表す波形図である。

【0073】このうち、図9は、有音無音判別器14にて算出されるフレーム最大電力  $P_f$ 、フレーム最大電力  $P_f$  と背景騒音電力  $P_b$  との差  $P_f - P_b$  と、音声符号化器12にて算出される1次及び2次の反射係数  $r_1$ 、 $r_2$  と、有音無音判別器14の出力である判定結果  $V$  とを表す波形図である。横軸は、時間で、縦軸は夫々符号付き16ビット整数表現した時の値である。

【0074】図に示すように、発声に基づく音声信号の

存在しない区間では、1次の反射係数  $r_1$  は、-1に近いほぼ一定値となり、また、2次の反射係数  $r_2$  は、+1に近いほぼ一定値となる。一方、音声信号の存在する区間では、反射係数  $r_1$ 、 $r_2$  は大きく変動し、音声信号の有無により顕著に異なる特性がよく表れている。

【0075】また、フレーム最大電力  $P_f$  と背景騒音電力  $P_b$  との差は、発声されていない区間では、比較的小さな値で変化し、発声の始端では、急激に値が増大する。このような特性を有するフレーム最大電力  $P_f$  と背景騒音電力  $P_b$  との差と、反射係数  $r_1$ 、 $r_2$  に基づいて、有音無音が判定された結果、発声に基づく音声信号の存在する区間は、確実に有音区間 ( $V=1$ ) と判定されている。

【0076】図10において、(a) は入力音声信号  $X_s$ 、(b) (c) は判定結果  $V$  と送信制御部16の制御状態とを表したものであり、特に (b) は本実施例による有音無音判別器14を用いた場合、(c) は1次及び2次の反射係数  $r_1$ 、 $r_2$  を用いることなく、フレーム最大電力  $P_f$  と背景雑音電力  $P_b$  との差が、第2しきい値  $TH_2$  以上であれば有音、第2しきい値  $TH_2$  より小さければ無音とする有音無音判定器を用いた場合である。

【0077】なお、送信状態の表示は、3が符号データを送信している状態、0が送信を停止している状態、1、2はプリアンプを送信している状態、-3、-2、-1はポストアンプを送信している状態に対応する。また、無音中に、周期的に送信されるポストアンプは、有音無音判定によるものではなく、先に説明したように、無音区間にて1000msに1回の割合で送信されるものである。

【0078】図10に示すように、フレームの電力値のみで判定を行っている (c) では、背景騒音だけが入力されている区間の数カ所で、有音区間と誤判定されているが、本実施例の (b) では、背景騒音だけが入力されている区間は、確実に無音区間と判定されている。

【0079】図11は、図10の2番目の発声の先頭部分を拡大したものである。図11に示すように、 $i$  フレームの途中から発声が始動されているのであるが、そのフレーム最大電力  $P_f$  は、サブフレーム毎に求められたサブフレーム電力の短期平均値  $P_{AS}$  の中から、最大のものが選択されている。このため、フレーム最大電力  $P_f$  としては、十分に大きな値が得られ、この  $i$  フレームは確実に有音フレームと判定される。

【0080】以上説明したように、本実施例においては、有音無音判別器14が、音声信号  $X_s$  が符号化される単位であるフレームを更に分割したサブフレーム毎にそのサブフレーム電力を算出し、更に、このサブフレーム毎の一つ前のサブフレーム電力との移動平均 (短期平均値)  $P_{AL}$  を算出し、同一フレームを構成するサブフレーム間でこの短期平均値  $P_{AL}$  を比較して、最大のものを

10

20

30

40

50



該フレームのフレーム最大電力 $P_f$ とし、このフレーム最大電力 $P_f$ に基づいて、該フレームの有音無音を判定するようにされている。

【0081】従って、本実施例によれば、発声がフレームの後半から開始されたとしても、そのフレームのフレーム最大電力 $P_f$ は、ほぼ音声信号の電力レベルに即したのものとなるため、このような場合にも、確実に有音フレームであると判断することができ、その結果、語頭が含まれたフレームの欠落による語頭切れを確実に防止でき、優れた通話品質を確保することができる。

【0082】また、本実施例によれば、フレーム最大電力 $P_f$ 及び背景騒音電力 $P_b$ をサブフレーム電力の移動平均により求めているので、突発的なノイズが重畳されたとしても、その影響を緩和することができ、ノイズによる誤判定を防止できる。また更に、各フレームの有音無音の判定を、そのフレーム最大電力 $P_f$ と背景騒音電力 $P_b$ との差だけによらず、周波数スペクトルの包絡形状の相違が反映される一次及び2次の反射係数 $r_1$ 、 $r_2$ を用いて行っているので、背景騒音が大きく、フレーム最大電力 $P_f$ と背景騒音電力 $P_b$ との差からでは判定が困難な場合であっても、確実に有音無音を判定できる。

【0083】そして、このように、フレームの有音無音を正確に判定できることにより、効果的にVOX制御を行うことができる。即ち、無音区間において、無音フレームを有音フレームとしてしまう誤判定が発生すると、直ちに有音区間に遷移して、送信制御部16により、送信器18が起動され、無駄な電力が消費されてしまうことになるが、このようなことを確実に防止できるのである。

【0084】また、本実施例では、有音区間において、無音フレームが検出されても、直ちに無音区間に変化するのではなく、所定フレーム以上連続して無音フレームを検出した場合に、無音区間に変化するようにされている。従って、本実施例によれば、会話中の息継ぎ等、短い無音状態が無音区間と判断されることがなく、有音区間/無音区間の変化時に実施されるVOX制御のための処理に基づく音声の欠落を防止することができる。即ち、本実施例のシステムでは、VOX制御によって送信器18を停止/起動することを、受信側に通知するために、有音区間から無音区間に変化した場合には3フレーム分のポストアンブルを、逆に無音区間から有音区間に変化した場合には2フレーム分のプリアンブルを送信するようにされているため、一旦、有音区間から無音区間に変化すると、その直後にまた有音区間に変化したとしても、ポストアンブル、プリアンブルの送信により、最低5フレーム(100ms)は、音声を送信することが

できず、その間に発生した音声は欠落してしまうのであるが、息継ぎのような短時間の無音状態は、これを無音区間として処理するのではなく、有音区間として処理することにより、このような音声の欠落を防止できるのである。

【0085】なお、上記実施例では、1次及び2次の反射係数を使用しているが、1次の反射係数のみ、或は2次の反射係数のみを使用したり、3次以上の反射係数を併用してもよい。また、上記実施例では、フレームが20msであるが、より長いフレーム(例えば40ms)を使用する場合に適用してもよく、この場合、サブフレームの大きさが同じであれば、同じ精度で語頭切れを防止できる。

#### 【図面の簡単な説明】

【図1】 本実施例の有音無音判別器の構成を表すブロック図である。

【図2】 本実施例の有音無音判別器が適用された自動車電話システムに用いられる移動局の送信部の構成を表すブロック図である。

【図3】 送信制御部にて生成される送信フレームの構成を表す説明図である。

【図4】 サブフレーム電力算出処理を表すフローチャートである。

【図5】 背景騒音電力、フレーム最大電力算出処理を表すフローチャートである。

【図6】 背景騒音電力、フレーム最大電力の説明図である。

【図7】 有音無音判定処理を表すフローチャートである。

【図8】 区間判定処理を説明するための状態遷移図である。

【図9】 有音無音判別器各部の信号波形を表す波形図である。

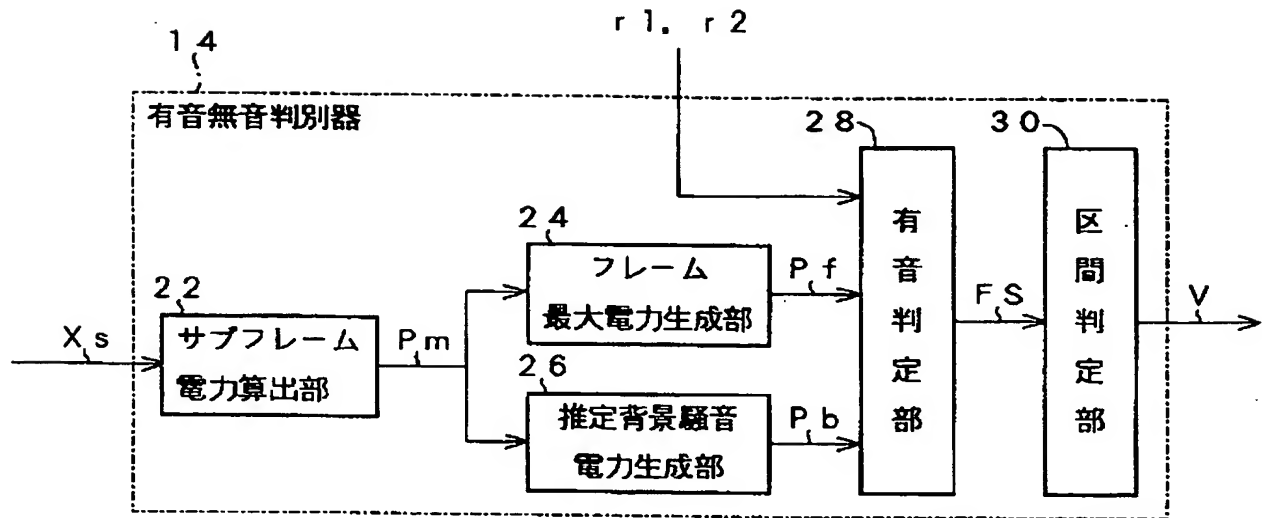
【図10】 有音無音判別器の判定結果、及び送信制御部の動作状態を表す説明図である。

【図11】 無音から有音への変化時におけるフレーム最大電力の算出結果を表す説明図である。

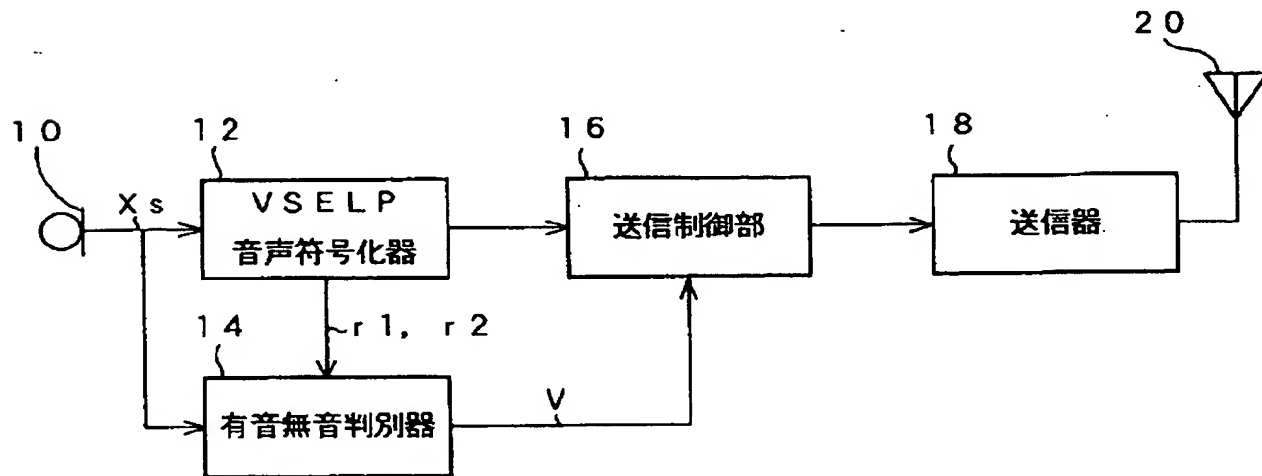
#### 【符号の説明】

10…入力部      12…音声符号化器      14…有音無音判別器  
16…送信制御部      18…送信器      20…アンテナ  
22…サブフレーム電力算出部      24…フレーム最大電力生成部  
26…推定背景騒音電力生成部      28…有音判定部  
30…区間判定部

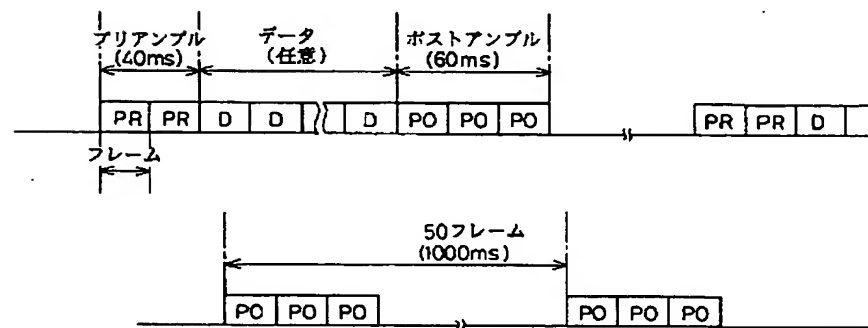
【図1】



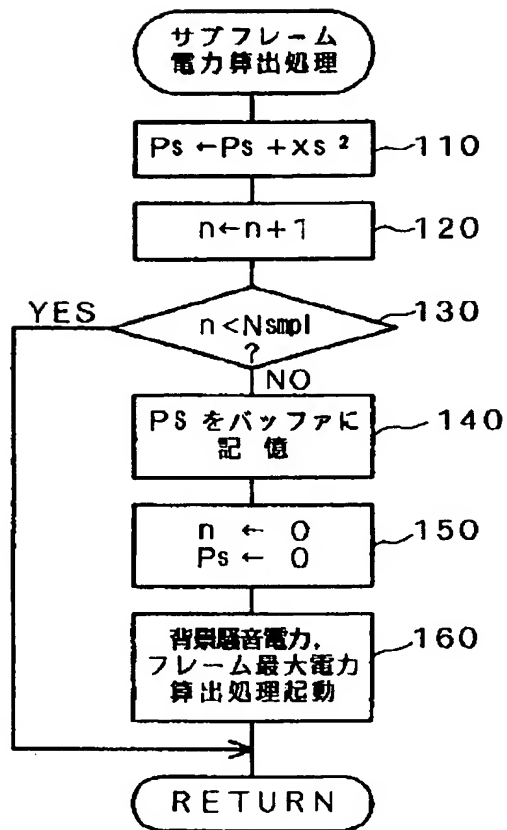
【図2】



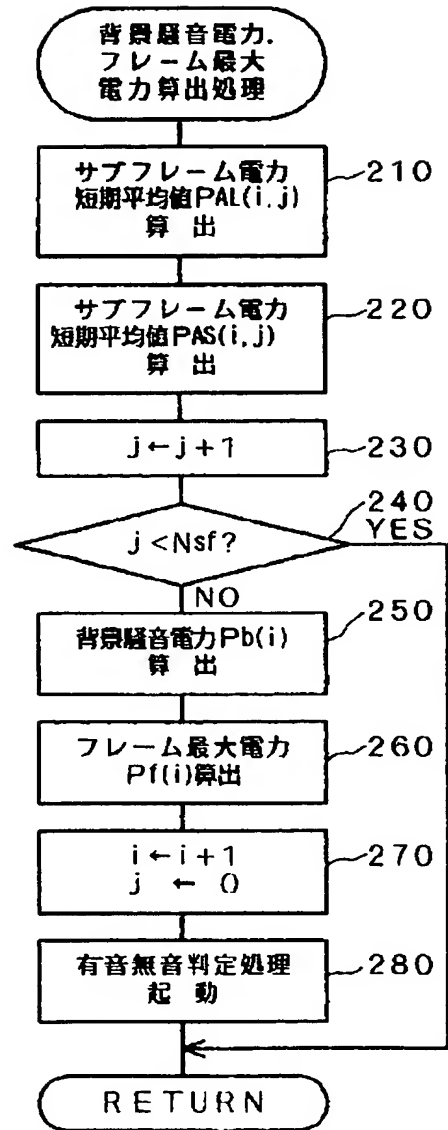
【図3】



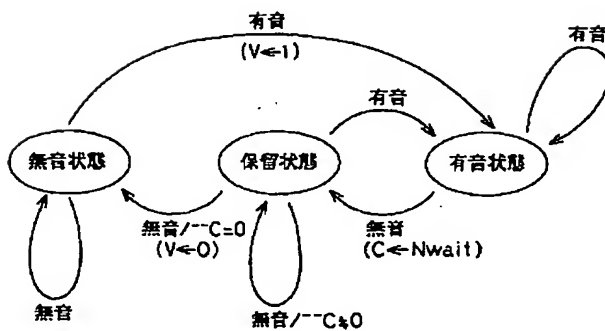
【図4】



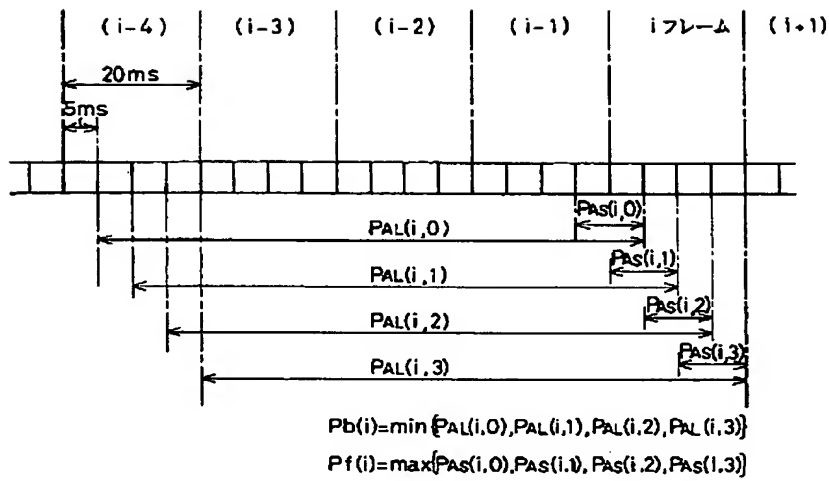
【図5】



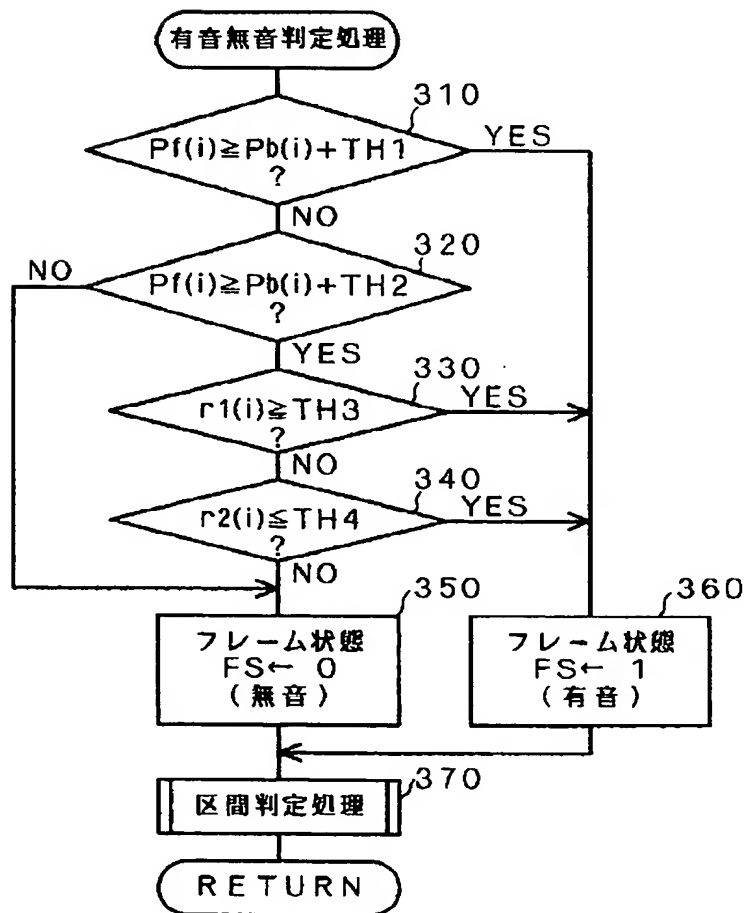
【図8】



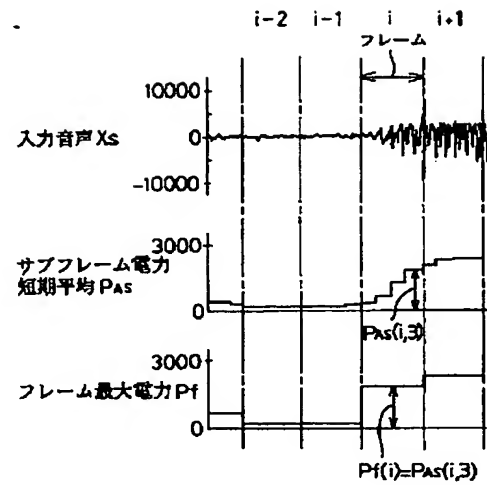
【図6】



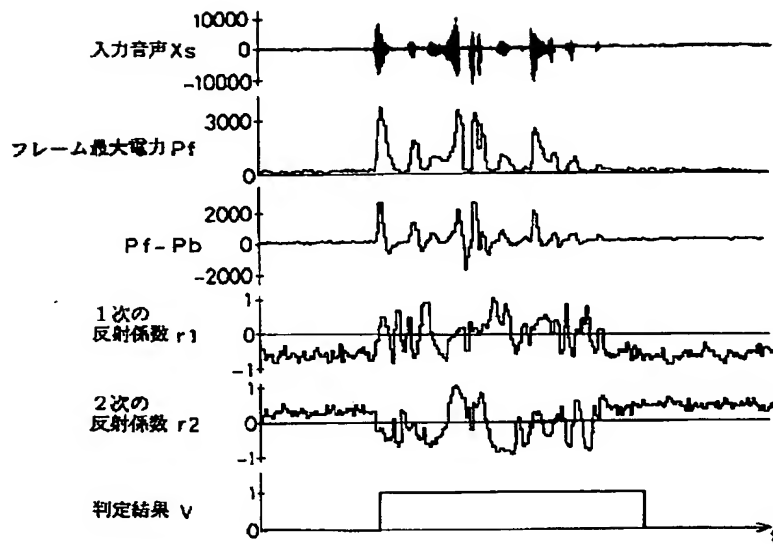
【図7】



【図11】



【図9】



【図10】

